AD-A058 424 CATHOLIC UNIV OF AMERICA WASHINGTON D C HUMAN PERFOR--ETC F/G 5/10
FEATURE EXTRACTION AND DECISION PROCESSES IN THE CLASSIFICATION--ETC(U)
JUL 78 J H HOWARD, J A BALLAS, D C BURGY N00014-75-C-0308
UNCLASSIFIED TR-78-4-ONR NL

| OF |
AD
A058424

END
DATE
FILMED
10-78
DDC

FEATURE EXTRACTION AND DECISION PROCESSES IN THE CLASSIFICATION OF AMPLITUDE MODULATED NOISE PATTERNS

James H. Howard, Jr., James A. Ballas, and Donald C. Burgy

ONR CONTRACT NUMBER N00014-75-C-0308

Technical Report ONR-78-4

Human Performance Laboratory
Department of Psychology
The Catholic University of America

July, 1978

Unclassified (14) TR-78-4-ONR

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ONR-78-4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) FEATURE EXTRACTION AND DECISION PROCESSES IN THE CLASSIFICATION OF AMPLITUDE MODULATED NOISE PATTERNS.		5. TYPE OF REPORT & PERIOD COVERED (9) Technical Rep[redacted]t,
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) James H. Howard, Jr., James A. Ballas[redacted] Donald C. Burgy		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-75-C-0308
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Catholic University of America Washington, D. C. 20064		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 197-027
11. CONTROLLING OFFICE NAME AND ADDRESS Engineering Psychology Programs, Code 455 Office of Naval Research		12. REPORT DATE (11) Jul[redacted] [redacted]78
		13. NUMBER OF PAGES (12) 69p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) auditory perception; auditory pattern recognition; auditory classification; feature extraction; decision processes; multidimensional scaling		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The relation between the perceptual features identified in a multidimensional scaling (MDS) analysis and the decision stage of the auditory classification process was investigated in four experiments based upon a set of sixteen complex acoustic patterns. The sounds consisted of broad-band white noise, amplitude modulated by sawtooth waves of varying frequency and attack. A psychological feature representation of the stimuli was obtained in Experiment 1 using a MDS analysis (INDSCAL) of the listeners'

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

409381

78 08 31 047

pairwise similarity ratings. Two groups of listeners in Experiment 2 learned to classify each of the sixteen signals into one of eight categories (two sounds per category). The two groups learned eight-category partitions that emphasized different features of the stimuli. Confusion matrices were analyzed in terms of both the stimulus space obtained in Experiment 1 and a probabilistic model of the listener's decision process. The model provided a reasonable fit to the observed data. Experiments 3 and 4 further tested the assumptions of the decision model. In Experiment 3, listeners were required to classify each member of a large set of amplitude modulated signals that formed a "grid" over the perceptual feature space. Subjective probability density functions for the eight categories estimated from listener responses using potential function or Parzen estimator techniques were consistent with those assumed by the model. In Experiment 4, MDS techniques were used to investigate the "conceptual space" underlying the listeners' memory for each of the eight categories in both groups. Category coordinates obtained from the MDS analysis corresponded well to the category centroids computed from the perceptual space of Experiment 1. Overall, results of the four experiments indicated that listeners employed an optimum-processor strategy to determine the relative importance of each feature in the decision process. The findings indicate that any theoretical treatment of auditory pattern recognition must address the interaction of the feature extraction and decision processes.

INTRODUCTION

Recent years have witnessed major theoretical advances in our understanding of the perceptual processes involved in the detection and discrimination of simple acoustic stimuli. In contrast, relatively little is known about the psychological processes that underlie the classification and recognition of complex acoustic patterns. A popular approach to the analysis of this problem assumes that human auditory recognition involves several distinct information-processing stages. A possible four-stage model of the auditory recognition process is diagrammed in Figure 1.

Insert Figure 1 here

According to this model, an unknown stimulus undergoes several transformations before it is recognized. First, an initial sensory representation of the signal is formed, and a preliminary analysis of the signal is completed. These processes are typically assumed to occur in the auditory periphery and have been reasonably well-specified in recent psychoacoustic research (e.g., Siebert, 1968; Dallos, 1973). Second, this preliminary "receptor" representation is further transformed or reorganized into a set of distinctive auditory features. This stage is referred to as feature extraction and is generally thought to involve the reduction of a stimulus to its essential characteristics (e.g., Anderson, Silverstein, Ritz & Jones, 1977). Third, this highly processed feature representation is

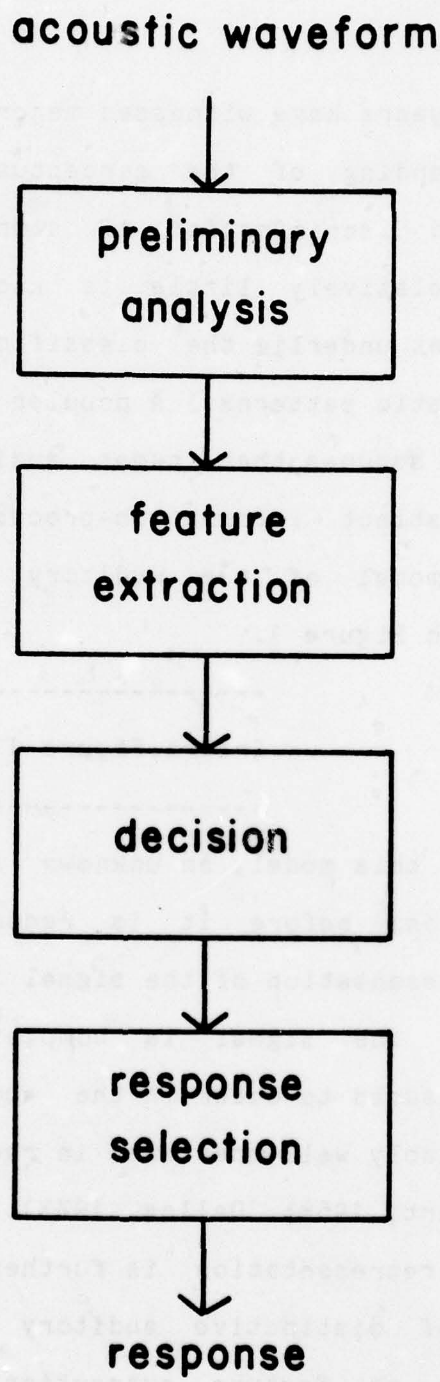

Figure 1. Flow diagram of a four stage pattern recognition model.

compared with information stored in memory to determine its classification and/or identify its structure (i.e., the relations among features). The processes involved in this stage may be extremely complex, and in the present model they are collectively referred to as the decision stage. Finally, an overt response may be initiated depending on the listener's task.

As suggested above, much psychoacoustic research has emphasized the basic psychophysical processes involved in pitch perception or the detection of pure tones, and their relation to underlying physiological functions (e.g., Evans & Wilson, 1977). As a result, a firm basis exists on which to speculate about the transduction and "preliminary analysis" stage of auditory pattern recognition. Unfortunately, in the case of complex acoustic patterns, no similar extensive empirical foundation exists on which to build a detailed model of the second (feature extraction) and third (decision) processing stages. The present paper focuses on the feature extraction stage and its relation to the decision process in an attempt to establish a firmer basis for a theoretical treatment of the auditory recognition problem.

Although no single theoretical statement of the feature extraction process exists, recent research has stressed its importance in auditory perception. As Anderson et al. (1977) have noted, "Distinctive features are usually viewed as a system for efficient preprocessing, whereby a noisy stimulus is reduced to its essential characteristics and decisions are made on these" (p. 429). In other words, the feature extraction process is "tuned" to select perceptually important information from the

output of the preliminary analysis stage, and discard information that is likely to be unimportant (Howard & Ballas, 1978). Since pattern recognition performance ultimately depends on the feature extraction process, a number of investigators have sought to specify those acoustic cues that are of primary psychological importance in the perception of complex acoustic patterns under various listening conditions.

The object of their investigation, the feature representation or output of the feature extraction stage, is obviously not directly observable and therefore must be inferred using indirect methods. Although a variety of techniques are available, multidimensional scaling has emerged as a useful method for identifying the underlying psychophysical structure of complex sounds (Plomp, 1976). Typically, listeners are asked to provide pairwise dissimilarity judgments on the set of signals of interest. A specific multidimensional scaling algorithm is then applied to decompose the resulting subjective proximity matrix into an n-dimensional metric space in which each signal is represented as a single point or vector. Although individual scaling methods vary widely in their underlying assumptions, Shepard (1972a) has noted that most are similar in that (1) they assume that the distances between stimuli in the underlying feature space are a monotonic function of the corresponding similarity judgments in the observed data, and (2) they employ an iterative procedure to obtain the perceptual space which best fits the observed data. A measure (e.g., "stress" in Kruskal, 1964) is frequently provided which reflects the degree of

discordance between the interstimulus distances in the n-dimensional stimulus space and the observed dissimilarity judgments.

Providing that a scaling solution with satisfactory stress exists, it is generally assumed that dimensions of the psychological stimulus space reflect those features that the listeners used to compare the stimuli. In interpreting the scaling solution, the investigator examines the relation between the perceptual space and the known physical structure of the stimuli. The outcome of this comparison can reveal the specific psychophysical transformations involved in the feature extraction process. These techniques have been used successfully to investigate the underlying psychological features involved in the perception of speech (Klein, Plomp & Pols, 1970; Shepard, 1972b), music-like sounds (Plomp & Stenneken, 1969; Miller & Carterette, 1975; Grey, 1977), and other complex non-speech sounds (Cermak & Cornillon, 1976; Howard & Silverman, 1976; Morgan, Woodhead & Webster, 1976; Howard, 1977).

Once the feature extraction process has transformed the stimulus into its essential characteristics, the decision process operates to classify or recognize the pattern. In the ideal case, the feature representation would unambiguously determine the true classification of a stimulus. In this case the task of the decision stage would be relatively straightforward. It need only partition the feature space into regions corresponding to the discriminable stimulus categories. Unfortunately, it is more likely the case that the output of the feature extraction process

is quite noisy and a considerably more complex decision process is called for. In particular, since the decision stage must operate in the presence of uncertainty, it can only evaluate the relative likelihood that a particular feature representation belongs to each category. Given this information, the decision processor may select the most likely source (i.e., category) for an unknown stimulus.

To this point, we have only considered the role of sensory information (i.e., the output of the feature extraction stage) in the decision process. As Green and Swets (1966) have pointed out in their elegant application of statistical decision theory to auditory detection, other utility or response-bias factors will also influence the decision process. These factors include the listener's estimate of the overall likelihood or a priori probability of specific categories as well as his consideration of the consequences of the decision. Although the decision process has been extensively investigated in the auditory detection situation, its role in the classification of complex auditory patterns has been neglected.

The overall question addressed in the present paper concerns the relation between the perceptual features identified in a multidimensional scaling analysis and the decision stage of the auditory classification process. In a classification task the listener is required to distinguish among a specified set of acoustic patterns. Consequently, one would expect the decision process to selectively emphasize one or another distinctive feature, depending on the configuration of stimuli in the

perceptual pattern space. For example, given a set of stimuli which differ in both pitch and loudness, listeners would likely use both features to evaluate pairwise similarity. On the other hand, if the same signals were then grouped into two categories based on only a single dimension (e.g., high and low pitch), then listeners learning this partition need only consider a single feature (i.e., pitch) to achieve optimal classification performance.

The present study investigates listener classification performance on a set of sixteen complex acoustic patterns. The signals consist of a broadband white noise carrier, amplitude modulated by sawtooth waves of varying frequency and attack. These signals were selected for investigation because of their similarity to a broad class of sounds frequently encountered in passive sonar environments (i.e., propeller cavitation). In Experiment 1 a multidimensional scaling analysis was performed on listeners' pairwise similarity ratings of the entire set of sixteen sounds. The primary purpose of this experiment was to obtain a psychological feature representation of the stimuli for use in the subsequent analyses. In Experiment 2, two groups of listeners learned to classify each of the sixteen signals into one of eight categories (two sounds per category). Each group learned a different category partition. The two eight-category partitions were selected to require the listeners to focus primarily on one of the two features (i.e., either modulation frequency or attack). The confusion matrices obtained in this experiment will be discussed in terms of both the perceptual

stimulus space identified in Experiment 1, and a probabilistic model of the listener's decision process. In Experiment 3, the same listeners were required to classify each member of a large set of amplitude modulated signals generated by factorially combining eleven values of attack and fifteen values of modulation frequency. A probability density function for each of the eight categories was estimated from listener responses using potential function or Parzen estimator techniques (e.g., Meisel, 1972). The results of this analysis will be compared with the findings of Experiment 2. Finally, in Experiment 4, multidimensional scaling techniques were used to investigate the "conceptual space" underlying the listeners' memory for each of the eight categories in both groups.

I. EXPERIMENT 1

The present experiment was designed to determine a precise, quantitative psychological feature representation of the sixteen amplitude modulated noise signals. Each listener was required to rate the pairwise similarity of all 120 possible pairs of the sixteen sounds. The INDSCAL multidimensional scaling program (Carroll & Chang, 1970) was used to determine a perceptual space for the signals. The INDSCAL model assumes that stimulus similarity is a decreasing linear function of the interstimulus distance in an underlying stimulus space. Unlike many metric scaling programs, the INDSCAL analysis produces both an overall normalized group stimulus space, and a vector of saliency weights for each listener reflecting the relative importance or salience of each dimension for that person. The group space reflects

those common features used by all or most listeners, and the saliency weights may be thought of as scaling factors to expand or contract each of the common dimensions for each observer. The INDSCAL model was used to evaluate feature consistency across individual listeners.

A. METHOD

1. Participants

Thirty student volunteers (twenty males, ten females) were paid \$9.00 to participate in the experiment. None of the listeners reported any history of hearing disorders.

2. Apparatus

All experimental events were controlled by a laboratory digital computer. The sawtooth modulation waveforms were synthesized by the computer and output on a 12 bit digital-to-analog converter (5 kHz sampling rate). This modulation signal was low-pass filtered (Krohn-Hite Model 3550, 2 kHz cutoff) and applied to the modulation input of a laboratory-constructed transconductance operational amplifier circuit (RCA CA3084). The carrier input to the operational amplifier was a 20 Hz - 20 kHz noise with a -3 dB/octave spectrum (B & K Type 1402 Random Noise Generator). The output gain of the transconductance operational amplifier circuit was directly proportional to the amplitude of the modulation signal. Hence, the circuit output consisted of amplitude modulated noise with an envelope determined by the modulation waveform characteristics (to be described below). This output signal was delivered to listeners over matched Telephonics TDH-49 headphones with

MX-41/AR cushions. The observers were isolated in a sound-attenuated booth throughout the experiment.

3. Stimuli

A set of sixteen amplitude modulated (100% modulation) white noise signals was constructed. In each case the signal envelopes were sawtooth functions varying in frequency and asymmetry. The modulation frequencies included 4, 5, 6, and 7 Hz, and the modulation waveform asymmetries included either 20 or 40 msec attack with gradual decay, or 40 or 20 msec decay with gradual attack. For example, a 4 Hz signal could have its maximum amplitude at 20, 40, 210, or 230 msec after the start of each period. An oscilloscope trace of two typical signals is displayed in Figure 2.

Insert Figure 2 here

Subjectively, the rapid attack signals have a "hammering" quality, whereas the gradual attack signals have a "sandpapering" quality. The signals were presented to the listeners at a comfortable listening level (64 dB SPL).

4. Procedure

Participants were seated individually in the sound-attenuated booth and heard instructions explaining their task. They were told that very dissimilar stimuli should be assigned a rating of "1", whereas very similar stimuli should receive a rating of "5". The remaining scale values were to be used for stimuli of intermediate similarity or dissimilarity.

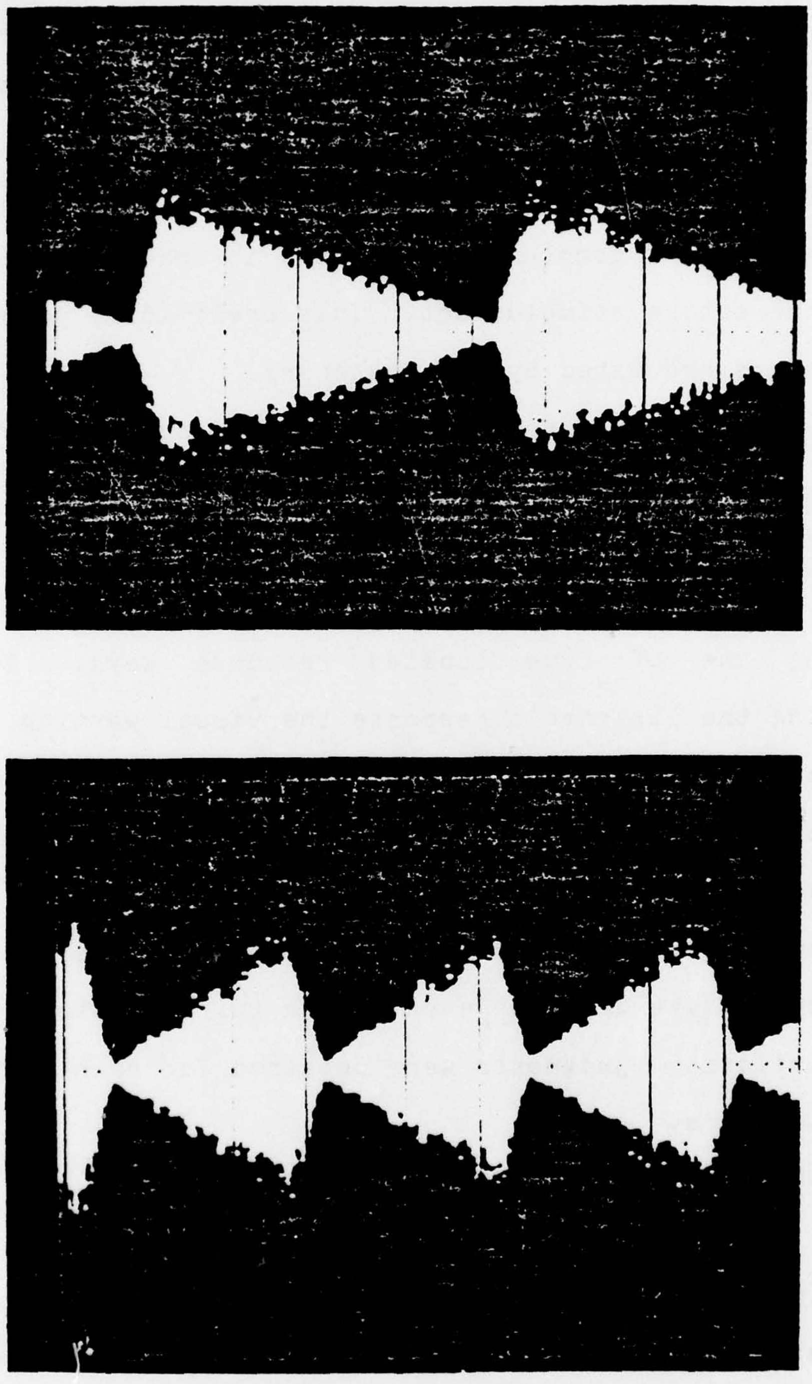

Figure 2. Oscilloscope trace of two typical amplitude modulated noise signals. The upper trace portrays a 4 Hz/20 msec attack signal, the lower a 7 Hz/20 msec decay signal.

The listeners were told to assign a similarity rating on the basis of their overall assessment of the stimulus similarity; no specific instructions were provided regarding the signal characteristics. Before beginning the experiment each listener heard a 3-second sample of each sound in order to become familiar with the entire stimulus set. This preliminary presentation was repeated as requested by the listeners.

Every trial began with a visual warning stimulus. After a short delay, a stimulus pair was presented successively in 3-sec segments with a 1-sec interstimulus interval.[1] After the stimuli were presented, the listener indicated the rated similarity by pressing one of five labeled response keys. Two seconds following the listener's response the visual warning occurred for the next stimulus pair. This procedure was repeated until each of the 120 possible pairs was presented twice, counterbalanced for order of presentation within trials. Signal pairs were presented in a random order. The above procedure was repeated on three successive days for each of the thirty listeners. In all, 180 similarity judgments were obtained for each of the possible stimulus pairs.

B. RESULTS AND DISCUSSION

A 16 by 16 off-diagonal asymmetric proximity matrix was determined for each listener and session by collapsing across the two similarity ratings for each signal pair within each session. The data from each of the three sessions for all thirty listeners were analyzed using the INDSCAL multidimensional scaling program. The resulting two dimensional scaling solution accounted for

approximately 69% of the overall variability. The normalized stimulus space for this solution is presented graphically in Figure 3.

Insert Figure 3 here

It is obvious from the geometric configuration of the stimuli that the two psychological dimensions or features correspond to the attack and modulation frequency parameters. In the following discussion the perceptual feature corresponding to attack will be referred to as signal Quality, and the feature corresponding to modulation frequency will be referred to as signal Tempo.[2]

A closer examination of Figure 3 suggests that Tempo bears a direct relation to modulation frequency, at least within the range of frequencies investigated. Further analysis substantiated this conclusion with 97.6% of the variability along this dimension being attributable to a linear function of modulation frequency ($T = .216M - 1.180$, where T designates Tempo, and M designates modulation frequency). In contrast, stimulus Quality appears to depend on both attack and modulation frequency since the stimuli tend to become somewhat closer together along this dimension as modulation frequency increases. Since the absolute duration of the attack/decay was held constant across modulation frequency, the proportion of each period spent in attack covaried with modulation frequency--percent attack increased with frequency for the rapid attack/gradual decay signals, and decreased with frequency for the gradual

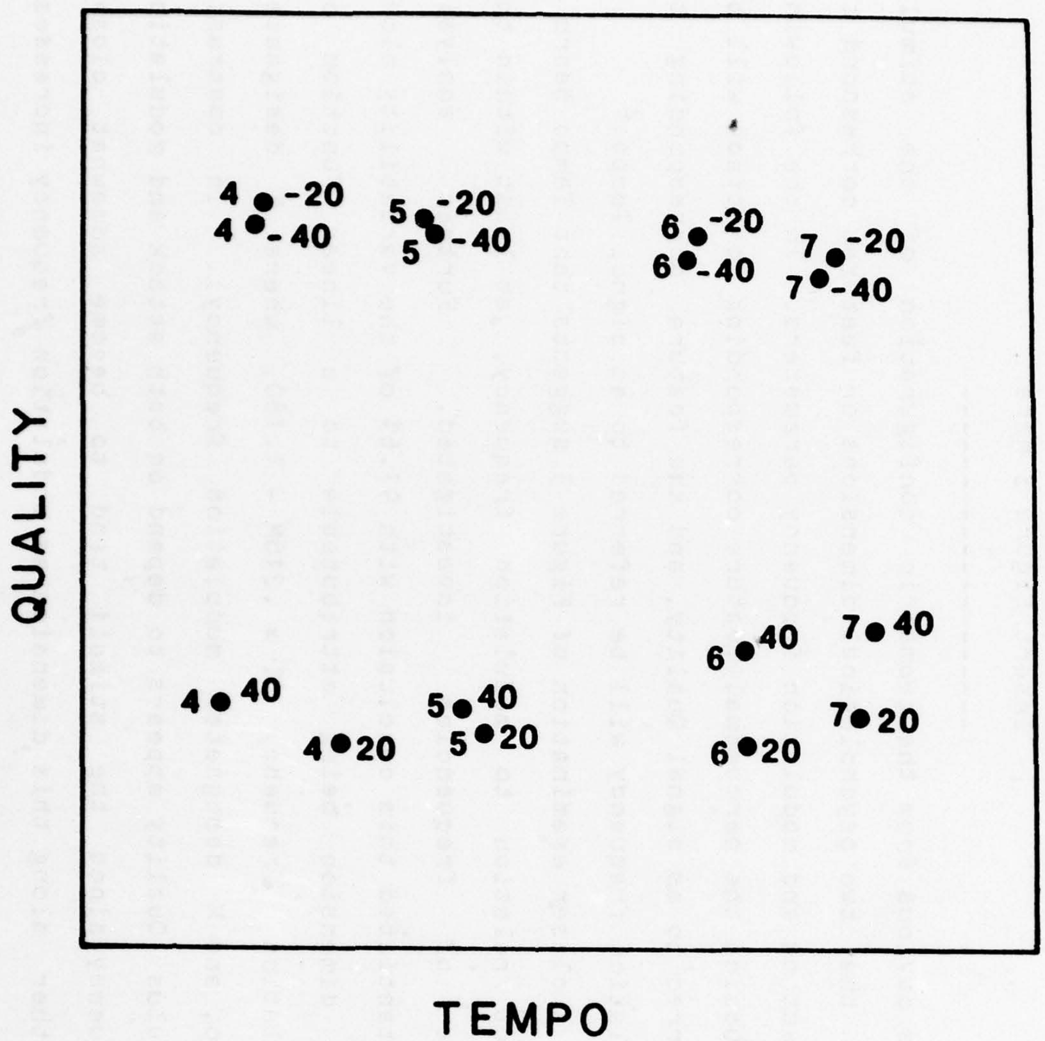

Figure 3. Normalized stimulus space for the two-dimensional INDSCAL solution, Experiment 1. The signal coordinates are presented in Table 1.

attack/rapid decay signals. It appears, therefore, that the relative amount of each period spent in attack, rather than the absolute duration of the attack is of primary psychological importance. This observation was confirmed statistically since signal Quality correlated more highly with the percent attack than it did with absolute attack duration ($r(15) = .994$ and $r(15) = .935$, respectively). Overall, 98.9% of the variance along the Quality dimension can be attributed to a linear function of percent attack ($Q = .007A - .364$, where Q refers to stimulus quality and A refers to percent attack).

As discussed above, a second outcome of the INDSCAL analysis is a weight vector for each individual listener that indicates the relative importance or salience of the two perceptual features. In the present data, 22 of the 30 listeners had larger saliency weights for the stimulus Quality dimension than for the stimulus Tempo dimension. This indicates that signal Quality was more important for these listeners than was signal Tempo. Overall, the Quality dimension accounted for considerably more of the variability in the inner-product matrix estimated from the judgment data (approximately 46%) than did the Tempo dimension (approximately 23%). This suggests that the "hammering" and "sandpapering" qualities of the stimuli were considerably more important in evaluating pairwise similarity than was the repetition rate.

II. EXPERIMENT 2

A. INTRODUCTION

The results of Experiment 1 have enabled us to characterize

precisely the perceptual feature representation of the sixteen stimuli. However, as indicated in the introduction, a question of primary interest concerns the relation between these features and subsequent processing stages in auditory pattern recognition. Specifically, we ask how the decision stage makes use of this information in determining a classification for the signal. In Experiment 2 we investigated this question by requiring two groups of listeners to learn different eight-category classifications of the sixteen stimuli. One of the two groups was required to distinguish two levels of Quality and four levels of Tempo in making their classification, whereas the other group discriminated two levels of Tempo and four levels of Quality. The specific question of interest concerns the possible relation between the classification partition learned and the feature information used by the decision process.

Clearly, the above empirical question may only be addressed in cases where a specific decision process has been specified. The following section outlines a simple probabilistic model of the decision stage. The model represents a generalization of previously proposed decision models for auditory signal detection (Green & Swets, 1966), pitch perception (Goldstein, 1973; Gerson & Goldstein, 1978), and visual recognition processes (Getty, Swets, Swets, & Green, in press).

As discussed in the general introduction, any theoretical treatment of the decision process must consider both sensory factors--resulting from sensory mechanisms--and utility or bias factors determined by nonsensory subjective task variables.

Although the following theory considers both factors, the primary focus of the present development is on the role of sensory factors in auditory recognition. In Experiment 2, variables traditionally thought to influence response bias (e.g., a priori category probability and response payoff) were held constant to minimize the importance of decision bias.

As indicated above, we assume that an initial preliminary analysis is performed on the incoming acoustic waveform to produce a vector of receptor measurements. This high-dimensional measurement vector, $\underline{m}$, is then transformed by an unspecified feature extraction processor, F, into a two-dimensional feature vector, $\underline{f}$, $F(\underline{m}) = \underline{f} = (f_T, f_Q)$. In the present context, we assume that the feature vector consists of two elements, Tempo and Quality. We assume further that moment-to-moment fluctuations or noise occurs in the outcome of the feature extraction process so that any specific presentation of a particular signal can result in any of a range of values for both Tempo and Quality. We assume that the feature values extracted for a particular sound are random variables sampled from Gaussian distributions with means equal to the "true" feature value, and standard deviations of σ_T and σ_Q for the Tempo and Quality dimensions, respectively.

After the stimulus has been analyzed into its feature vector, another transformation is applied by the decision processor to determine its classification. That is, $D(\underline{f}) = c^{(i)}$, where $c^{(i)}$ indicates that the signal has been assigned to category i. We assume that the decision processor operates by comparing

the feature representation of the unknown signal, $\underline{f}$, to a prototype or "ideal" representation for each of the eight categories. The listener's decision is then based on the likelihood that the signal occurred given each of the eight categories.[3] This, in turn, depends on the unknown signal's proximity to the prototype (i.e., the centroid) of each category in the perceptual feature space. In other words, the decision processor estimates the probability that the unknown signal occurred given each category, $Pr(\underline{f} \mid c^{(i)})$, i = 1,2,...,8.

Since uncertainty exists in the feature extraction process, the decision processor must estimate the precise location of each category prototype in the feature space. Further, since the features extracted for a particular sound are assumed to be orthogonal Gaussian random variables, the likelihood function for each category over the feature space is bivariate Gaussian with zero covariance. The likelihood function will have an identical shape for each of the categories, and will be centered at the category prototype. Therefore, the likelihood that a particular signal occurred given category $c^{(i)}$ is determined by

$$Pr(\underline{f} \mid c^{(i)}) = \frac{1}{2\pi \, |\underline{V}|^{1/2}} \exp\left[-1/2\,(\underline{f} - \underline{p}^{(i)})\,\underline{V}^{-1}\,(\underline{f} - \underline{p}^{(i)})'\right] \qquad [1]$$

where $\underline{p}^{(i)}$ is the prototype vector for category $c^{(i)}$ obtained by averaging the feature values across the two members of the category, $\underline{P}^{(i)} = (\underline{f}_1^{(i)} + \underline{f}_2^{(i)})/2$, $\underline{f}_1^{(i)}, \underline{f}_2^{(i)} \in c^{(i)}$, and $\underline{V}$ is the covariance matrix. Since in the present context the two features are assumed to be orthogonal, this matrix consists of variances for the Tempo and Quality features on the main diagonal

and zero elements elsewhere. $|\underline{V}|^{1/2}$ denotes the square root of the determinant of $\underline{V}$, in this case $\sigma_T \sigma_Q$, and $\underline{V}^{-1}$ indicates the inverse of $\underline{V}$.

An important assumption of the present model is that the listener's uncertainty regarding the two perceptual features can be reduced with experience in the classification task. In other words, listeners can "fine-tune" their feature extraction process to reduce the uncertainty or variability associated with a particular feature. Obviously this decrease occurs with a lower bound being determined by the absolute discriminability of each feature. More importantly, we assume that this reduction in variability with experience is under listener control, and that the listener can selectively adjust his or her variability on the two dimensions independently. In learning to classify a set of stimuli, observers can choose to focus their attention on one or another dimension and thereby reduce their uncertainty with respect to that dimension.

It should be clear that differences in the variance parameters influence the relative importance of the two features in the decision process, and hence determine classification performance. The lower the relative variance along a particular dimension in the feature space, the greater the effect of that feature in determining signal likelihood.[4] Therefore, one would expect the standard deviation parameters determined by the listener's selective attentional mechanisms to be based on the classification requirements of the task. Specifically, in the present experiment one would expect listeners in the two groups

to differentially emphasize signal Tempo and Quality depending on the category partition they are required to learn.

The results of the present classification experiment are examined in terms of the above model. In particular, the model is fit to individual listener confusion matrices by estimating the two variance parameters in Equation 1. The confusion matrices provide an estimate of the subjective a posteriori probabilities, i.e., the probability of category $c^{(i)}$ given a particular stimulus, $Pr(c^{(i)} | \underline{f})$. These estimated a posteriori probabilities can be compared to theoretical a posteriori probabilities derived using Bayes' rule

$$Pr(c^{(i)} | \underline{f}) = \frac{Pr(\underline{f} | c^{(i)})\ Pr(c^{(i)})}{\sum_{j=1}^{8} Pr(\underline{f} | c^{(j)})\ Pr(c^{(j)})} \qquad [2]$$

where, $Pr(c^{(i)})$ denotes the a priori probability of category $c^{(i)}$, which in the present case is assumed to be constant across categories ($Pr(c^{(i)}) = 1/8$ for all i) and $Pr(\underline{f} | c^{(i)})$ is obtained from Equation 1.[5] The two standard deviations in Equation 1, σ_T and σ_Q are then estimated by minimizing the sum of squared deviations between the theoretical and estimated probabilities using a standard gradient technique.

B. METHOD

1. Participants

Eight experimentally naive student volunteers were paid to participate in the experiment. Four (2 males, 2 females) served in the Tempo group, and four (2 males, 2 females) served in the Quality group. None of the listeners reported any history of

hearing disorders.

2. Apparatus

Same as Experiment 1.

3. Stimuli

Two eight-category partitions of the sixteen signals used in Experiment 1 were formed. One partition, presented to the Tempo group, emphasized stimulus Tempo by requiring listeners to discriminate four levels of Tempo and two of Quality. The second partition was presented to the Quality group and required four levels of Quality discrimination, and two of Tempo discrimination. Table 1 indicates the assignment of the sixteen signals to the eight categories for both groups.

Insert Table 1 here

----

4. Procedure

Listeners were tested individually in a sound-attenuated booth. They were told that their task was to learn to classify two sounds into each of eight categories, and that every sound they heard would correctly belong in only one category. No specific instructions were provided regarding how signal Tempo and Quality were to be used. Each trial began with a visual warning followed by a 3-sec presentation of one of the sixteen sounds. After the signal terminated, the listener depressed one of eight response keys (labeled 1-8) to indicate the category decision. Feedback was provided after each trial.

All listeners received 80 trials in each of nine sessions

Table 1. Perceptual signal coordinates and category assignments for both groups, Experiment 2.

	COORDINATES		CATEGORY	
SIGNAL	TEMPO	QUALITY	TEMPO GROUP	QUALITY GROUP
1	-.338	-.286	1	1
2	-.366	-.243	1	2
3	-.336	.271	2	3
4	-.325	.282	2	4
5	-.087	-.277	3	1
6	-.109	-.249	3	2
7	-.144	.257	4	3
8	-.145	.263	4	4
9	.190	-.285	5	5
10	.192	-.216	5	6
11	.127	.229	6	7
12	.142	.250	6	8
13	.315	-.259	7	5
14	.328	-.170	7	6
15	.274	.206	8	7
16	.284	.227	8	8

over three consecutive days for a total of 720 trials. Each of the sixteen signals was presented equally often in a random order.

C. RESULTS

1. Overall performance analysis

Overall performance was assessed by computing mean percent correct on each of the sixteen stimuli for each listener, collapsed across the three sessions within each day. The results of this analysis, further collapsed across stimuli, are presented in Figure 4 for the two groups. Several aspects of these data are of interest.

Insert Figure 4 here

First, in terms of overall responding, both groups are well above the chance level of 12.5%. By day 3, the very worst listener, ML in the Quality group, was responding at approximately four times the rate expected by chance alone.

Second, both the Tempo and Quality groups tended to show higher performance on days 2 and 3 than on day 1. Mean percent correct collapsed across the four listeners was 55, 75 and 75% on days 1, 2, and 3, respectively for the Tempo group, and 33, 47 and 51% on days 1, 2, and 3, respectively for the Quality group. This finding was confirmed statistically by a significant main effect of Day in a two-way (Group by Day) analysis of variance with repeated measures on the Day factor, $\underline{F}(2,12) = 35.94$, $\underline{p} < .001$.

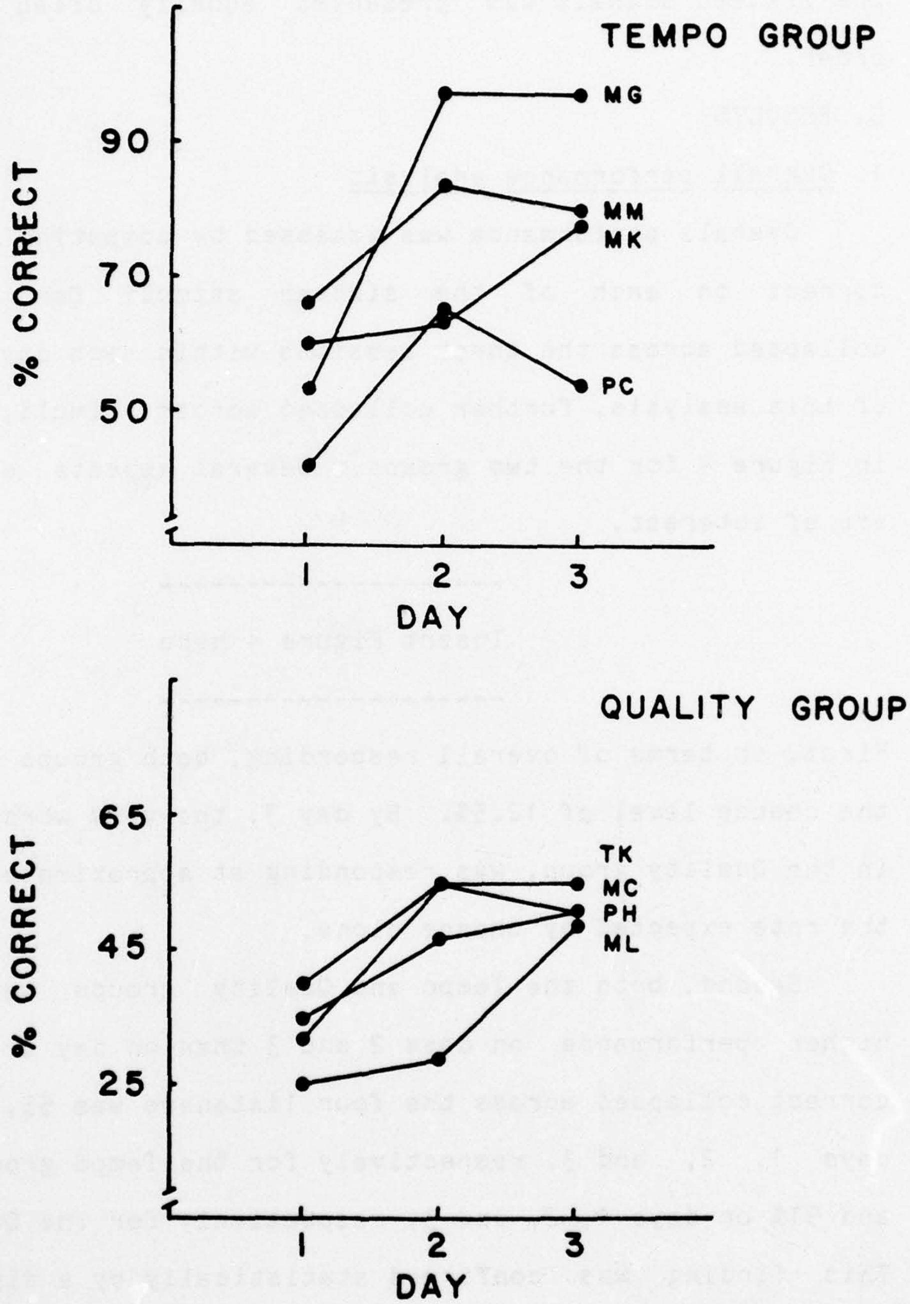

Figure 4. Mean percent correct overall by day and listener, Experiment 2.

Third, the Tempo group performed at a considerably higher level than did the Quality group (mean performance was 68 and 44% for the two groups, respectively). This observation was also supported statistically in the above analysis, $\underline{F}(1,6) = 15.55$, $\underline{p} < .01$. This finding indicates that the category partition learned by the Quality group was considerably more difficult than that learned by the Tempo group.

Another aspect of the performance data of potential interest is the percent correct observed for each of the sixteen stimuli. Table 2 displays mean day 3 performance data for each signal and listener in the experiment.

Insert Table 2 here

Examination of this table reveals that by day 3, all of the listeners in the Tempo group, and two of the listeners in the Quality group were classifying all stimuli at an above-chance level. The two exceptions to this, listeners PH and TK in the Quality group, classified three and two of the sixteen stimuli, respectively, at a chance or below-chance level. The only consistent trend observed across all listeners is an "anchoring" effect noted for signals occupying corner positions in the stimulus space. For the Tempo group, the four signals having extreme values on both features (i.e., signals 1, 4, 13 and 16) were more frequently correct than were the four signals having extreme values on neither feature (i.e., signals 6, 7, 10 and 11). Performance on the "corner" signals was 82%, whereas

Table 2. Percent correct for individual stimuli and listeners, day 3.

SUBJECT	SIGNAL 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
TEMPO GROUP																
MM	87	87	100	80	53	87	80	87	53	60	80	87	87	80	80	100
MG	93	87	93	100	73	93	87	93	73	67	87	100	93	80	93	80
PC	67	73	73	40	40	53	53	73	27	53	40	47	73	67	20	60
MK	73	93	93	93	47	67	73	67	67	60	73	93	100	100	80	73
MEAN	80	85	90	78	53	75	73	80	55	60	70	82	88	82	68	78
QUALITY GROUP																
PH	07	80	93	13	47	73	27	53	73	27	33	27	87	93	07	73
TK	40	73	00	73	40	60	40	33	80	93	07	73	87	100	33	53
ML	67	47	93	20	47	67	27	40	73	40	33	27	53	53	53	47
MC	67	53	60	60	27	40	27	47	60	47	40	40	60	87	33	60
MEAN	45	63	62	42	40	60	30	43	72	52	28	42	72	83	32	58

performance on the "inner" signals was 70%, $t(15) = 3.68$, $p < .005$. A similar, but statistically nonsignificant, trend was observed for the Quality group (54 and 43% for the corner and inner stimuli, respectively), $t(15) = 1.52$, $p > .05$. This finding is consistent with an end-anchoring effect noted in a variety of learning contexts.

2. Confusion matrix analysis

Although the overall performance data reported above are clearly important, a detailed analysis of the kinds of errors that listeners make is of primary importance in the present paper. A 16 by 8 (signal by category) confusion matrix was determined for each listener on each day by collapsing across the three sessions within each day. These 24 matrices (eight listeners by three days) formed the basis of all subsequent analyses.

Equations 1 and 2 were used to estimate a theoretical confusion matrix for each of the observed matrices. The theoretical matrices were determined by selecting standard deviation parameters ($\hat{\sigma}_T$ and $\hat{\sigma}_Q$, Equation 1) that minimized the discrepancy between the theoretical and observed matrices in a least squares sense. A standard, quasi-Newton gradient algorithm was used to perform the fits (subroutine ZXMIN in the IMSL statistical library). Fits were obtained from several starting points in the ($\hat{\sigma}_T$, $\hat{\sigma}_Q$) parameter space for randomly selected matrices as a precaution against unstable solutions resulting from local minima. Several outcomes of this analysis are discussed in detail below.

First, the model provided a reasonable fit to the observed confusion matrices under most conditions. Pearson product-moment correlation coefficients were computed between the theoretical and observed data for each of the matrices as a measure of goodness-of-fit. The results of this analysis are displayed in Table 3.

Insert Table 3 here

The theoretical matrix accounted for between 61 and 96% of the variance in the observed data in all but one case (listener ML, day 1, $\underline{r}^2$ = 40%). On the average, the model accounted for 82% of the variability for the Tempo group and 69% of the variability for the Quality group (72% if ML, day 1 is excluded). It should be noted that although many confusions never occur (i.e., some cells of the matrix are almost always zero), the present fits were obtained with only two free parameters and 128 estimated points. Sample theoretical and observed confusion matrices are presented for four representative day 3 cases in Table 4.

Insert Table 4 here

These data represent the best and worst fitting conditions for the Tempo (listeners MG and PC, respectively) and Quality (listeners MC and PH, respectively) groups.

Second, the standard deviation parameters estimated from the present data, $\hat{\sigma}_T$ and $\hat{\sigma}_Q$, are consistent with the assumption

Table 3. Pearson product-moment correlation coefficients computed between observed confusion matrices and the best-fitting theoretical matrices.

	day 1	day 2	day 3
TEMPO GROUP			
MM	.87	.97	.96
MG	.83	.98	.98
PC	.80	.92	.83
MK	.87	.88	.93
QUALITY GROUP			
PH	.78	.80	.84
TK	.81	.86	.88
ML	.63	.79	.89
MC	.78	.93	.95

Table 4. Observed and theoretical day 3 a posteriori probability matrices for the best fitting and worst fitting listeners in each group.

TEMPO GROUP, LISTENER MG (r = .98)

SIGNAL	CATEGORY 1		2		3		4		5		6		7		8	
	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.
1	.93	.98	.00	.01	.07	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
2	.87	.98	.00	.02	.13	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
3	.00	.02	.93	.97	.00	.00	.07	.01	.00	.00	.00	.00	.00	.00	.00	.00
4	.00	.02	1.00	.96	.00	.00	.00	.02	.00	.00	.00	.00	.00	.00	.00	.00
5	.00	.00	.00	.00	.73	.99	.00	.01	.20	.00	.07	.00	.00	.00	.00	.00
6	.07	.00	.00	.00	.93	.97	.00	.03	.00	.00	.00	.00	.00	.00	.00	.00
7	.00	.00	.00	.02	.00	.02	.87	.96	.00	.00	.07	.00	.07	.00	.00	.00
8	.00	.00	.07	.02	.00	.02	.93	.96	.00	.00	.00	.00	.00	.00	.00	.00
9	.00	.00	.00	.00	.20	.00	.00	.00	.73	.86	.07	.02	.00	.11	.00	.01
10	.07	.00	.00	.00	.13	.00	.00	.00	.67	.82	.00	.03	.13	.12	.00	.03
11	.00	.00	.00	.00	.00	.00	.13	.00	.00	.03	.87	.91	.00	.00	.00	.06
12	.00	.00	.00	.00	.00	.00	.00	.00	.00	.02	1.00	.88	.00	.00	.00	.10
13	.00	.00	.00	.00	.00	.00	.00	.00	.07	.14	.00	.00	.93	.82	.00	.03
14	.00	.00	.00	.00	.00	.00	.00	.00	.20	.09	.00	.00	.80	.82	.00	.09
15	.00	.00	.00	.00	.00	.00	.00	.00	.00	.02	.07	.09	.00	.06	.93	.83
16	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01	.20	.07	.00	.05	.80	.87

Table 4--Continued.

TEMPO GROUP, LISTENER PC (r = .83)

SIGNAL	CATEGORY 1		2		3		4		5		6		7		8	
	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.
1	.67	.65	.20	.15	.13	.14	.00	.06	.00	.00	.00	.00	.00	.00	.00	.00
2	.73	.67	.00	.18	.27	.10	.00	.05	.00	.00	.00	.00	.00	.00	.00	.00
3	.20	.15	.73	.59	.00	.03	.00	.22	.00	.00	.07	.00	.00	.00	.00	.00
4	.20	.14	.40	.58	.07	.04	.33	.24	.00	.00	.00	.00	.00	.00	.00	.00
5	.20	.09	.07	.03	.40	.60	.07	.14	.13	.08	.13	.05	.00	.01	.00	.01
6	.33	.12	.07	.04	.53	.58	.00	.16	.07	.05	.00	.04	.00	.00	.00	.00
7	.00	.05	.07	.21	.07	.14	.53	.53	.13	.01	.20	.07	.00	.00	.00	.00
8	.00	.05	.07	.21	.00	.13	.73	.53	.13	.01	.07	.07	.00	.00	.00	.00
9	.00	.00	.00	.00	.07	.05	.07	.01	.27	.45	.07	.11	.53	.28	.00	.11
10	.07	.00	.07	.00	.00	.04	.07	.01	.53	.41	.13	.14	.13	.26	.00	.14
11	.00	.00	.00	.00	.00	.04	.33	.06	.27	.14	.40	.45	.00	.06	.00	.24
12	.00	.00	.00	.00	.00	.03	.07	.05	.33	.13	.47	.45	.07	.07	.07	.27
13	.00	.00	.00	.00	.00	.01	.00	.00	.00	.31	.00	.06	.73	.47	.27	.15
14	.00	.00	.00	.00	.00	.00	.00	.00	.20	.27	.07	.08	.67	.45	.07	.21
15	.00	.00	.00	.00	.00	.00	.00	.00	.00	.13	.13	.26	.67	.17	.20	.43
16	.00	.00	.00	.00	.00	.00	.00	.00	.27	.12	.13	.25	.00	.17	.60	.45

Table 4--Continued.

QUALITY GROUP, LISTENER MC ($\underline{r}$ = .95)

SIGNAL	CATEGORY 1		2		3		4		5		6		7		8	
	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.
1	.67	.53	.33	.45	.00	.00	.00	.00	.00	.02	.00	.00	.00	.00	.00	.00
2	.47	.43	.53	.55	.00	.00	.00	.00	.00	.01	.00	.01	.00	.00	.00	.00
3	.00	.00	.00	.00	.60	.48	.40	.48	.00	.00	.00	.00	.00	.02	.00	.02
4	.00	.00	.00	.00	.40	.48	.60	.49	.00	.00	.00	.00	.00	.02	.00	.02
5	.27	.43	.40	.36	.00	.00	.00	.00	.33	.15	.00	.06	.00	.00	.00	.00
6	.27	.38	.40	.42	.00	.00	.00	.00	.20	.12	.13	.08	.00	.00	.00	.00
7	.00	.00	.00	.00	.27	.39	.67	.39	.00	.00	.00	.00	.07	.10	.00	.11
8	.00	.00	.00	.00	.27	.40	.47	.40	.00	.00	.00	.00	.27	.10	.00	.11
9	.07	.12	.07	.08	.00	.00	.00	.00	.60	.60	.27	.20	.00	.00	.00	.00
10	.00	.06	.00	.07	.00	.00	.00	.00	.47	.36	.47	.50	.07	.00	.00	.00
11	.00	.00	.00	.00	.13	.09	.20	.09	.00	.00	.00	.00	.40	.41	.27	.41
12	.00	.00	.00	.00	.00	.10	.20	.10	.00	.00	.00	.00	.40	.38	.40	.42
13	.00	.03	.00	.03	.00	.00	.00	.00	.60	.59	.40	.35	.00	.00	.00	.00
14	.00	.01	.00	.01	.00	.00	.00	.00	.13	.21	.87	.77	.00	.00	.00	.00
15	.00	.00	.00	.00	.00	.02	.00	.02	.00	.00	.00	.00	.33	.50	.67	.45
16	.00	.00	.00	.00	.00	.03	.00	.02	.00	.00	.00	.00	.40	.47	.60	.48

Table 4--Continued.

QUALITY GROUP, LISTENER PH ($\underline{r}$ = .84)

SIGNAL	CATEGORY 1		2		3		4		5		6		7		8	
	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.	OBS.	THE.
1	.07	.38	.87	.62	.00	.00	.07	.00	.00	.00	.00	.00	.00	.00	.00	.00
2	.20	.23	.80	.77	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
3	.00	.00	.00	.00	.93	.53	.07	.47	.00	.00	.00	.00	.00	.00	.00	.00
4	.00	.00	.00	.00	.87	.52	.13	.48	.00	.00	.00	.00	.00	.00	.00	.00
5	.47	.75	.53	.25	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
6	.27	.65	.73	.35	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
7	.00	.00	.00	.00	.27	.48	.60	.52	.00	.00	.00	.00	.13	.00	.00	.00
8	.00	.00	.00	.00	.27	.47	.53	.53	.00	.00	.00	.00	.13	.00	.07	.00
9	.00	.00	.20	.00	.00	.00	.00	.00	.73	.83	.07	.17	.00	.00	.00	.00
10	.27	.00	.00	.00	.00	.00	.00	.00	.40	.42	.27	.58	.00	.00	.07	.00
11	.00	.00	.00	.00	.07	.00	.20	.00	.00	.00	.00	.00	.33	.57	.40	.43
12	.00	.00	.00	.00	.00	.00	.33	.00	.00	.00	.00	.00	.40	.51	.27	.49
13	.00	.00	.00	.00	.00	.00	.00	.00	.87	.65	.13	.35	.00	.00	.00	.00
14	.00	.00	.00	.00	.00	.00	.00	.00	.07	.13	.93	.87	.00	.00	.00	.00
15	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.07	.48	.93	.52
16	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.27	.44	.73	.56

that, with experience, listeners can selectively and independently adjust their uncertainty on the two perceptual features. The estimated parameters are displayed in Table 5 for all conditions.

Insert Table 5 here

It is evident from this table that, in general, uncertainty decreases over days in the experiment. Since the parameters are estimated by fitting distributions to observed confusion matrices, it is not surprising that uncertainty decreases as performance improves. What is more significant is the observation that the two parameters are dramatically different for the two groups. In particular, for all listeners in the Tempo group, $\hat{\sigma}_T$ was substantially smaller than $\hat{\sigma}_Q$ (overall means of .101 and .227, respectively). In contrast, the Quality group showed less Quality uncertainty than Tempo uncertainty (overall means of .110 and .290, respectively), and by day 3 all listeners in the Quality group had a lower $\hat{\sigma}_Q$ than $\hat{\sigma}_T$. Since the magnitude of these parameters is inversely related to the relative importance of their corresponding features in the decision process, this finding indicates that signal Tempo was given a greater emphasis by the Tempo group, whereas signal Quality was given greater emphasis by the Quality group.

Of further interest is the finding that on day 1 all listeners in the Tempo group were emphasizing Tempo relative to Quality, while only two of the listeners in the Quality group (PH

Table 5. Estimated standard deviation parameters for both features and all conditions, Experiment 2.

TEMPO GROUP	day 1		day 2		day 3	
	$\hat{\sigma}_T$	$\hat{\sigma}_Q$	$\hat{\sigma}_T$	$\hat{\sigma}_Q$	$\hat{\sigma}_T$	$\hat{\sigma}_Q$
MM	.108	.283	.073	.178	.076	.208
MG	.107	.446	.064	.179	.066	.192
PC	.130	.649	.112	.276	.137	.324
MK	.134	.275	.155	.217	.063	.248
MEAN	.120	.413	.101	.213	.083	.243
QUALITY GROUP						
PH	.373	.208	.246	.070	.061	.054
TK	.244	.208	.220	.033	.205	.034
ML	.276	.424	.221	.262	.217	.074
MC	.235	.299	.208	.044	.220	.062
MEAN	.282	.285	.224	.012	.176	.056

and TK) revealed an analogous emphasis on signal Quality. The other listeners in this group (ML and MC) emphasized Tempo early in the experiment, and for one of these listeners, ML, this trend did not reverse until the last day of the experiment.

D. DISCUSSION

It is clear from these results that the decision model outlined above provides a reasonable description of how feature information is used by the decision processor in an auditory classification task. The findings are also clear in supporting the specific assumption that listeners can selectively and independently adjust the relative importance of the two perceptual features. However, despite this consistency, two major questions remain unanswered. First, at present it is unclear how the uncertainty parameters estimated from the data relate to the listener's sensitivity to the attack and modulation frequency cues. What do the values obtained for these parameters mean in terms of listener sensitivity? Second, although it is intuitively reasonable to argue that listeners in the Tempo group should stress Tempo relative to Quality, and that listeners in the Quality group should stress Quality relative to Tempo, it is not clear why they select the specific values observed. What criterion does the listener use to determine the importance of one feature relative to another? Both issues are considered further below.

Consider the relation between the specific value of each standard deviation parameter and listener sensitivity to the corresponding feature. By determining the separation between

individual category prototypes (i.e., centroids) along each of the two dimensions separately, we may examine the relation between the $\hat{\sigma}_T$ and $\hat{\sigma}_Q$ parameters and performance for the two groups. In other words, the standard deviation parameters give us an idea of the extent of overlap between adjacent likelihood functions in the perceptual space.

Figure 5 displays a hypothetical perceptual space for the Tempo and Quality features showing representative (i.e., median) intercategory distances and mean day 3 one-standard-deviation contours for the two groups.

Insert Figure 5 here

First, it should be noted that the typical intercategory separation along the Tempo dimension is considerably smaller for the Tempo group than for the Quality group (.220 vs .457). In contrast, the median separation along the Quality dimension is substantially greater for the Tempo group than for the Quality group (.506 vs .058). Second, it is also obvious from the figure that the standard deviation parameters for the two groups parallel the median separations. The smallest mean standard deviations correspond to the smallest intercategory distances. Although it appears that listeners in the Tempo group were able to adjust their standard deviations along the two dimensions to produce relatively little overlap in the likelihood functions, considerable overlap exists along the Quality dimension for the Quality group. It seems that the listeners were not able to

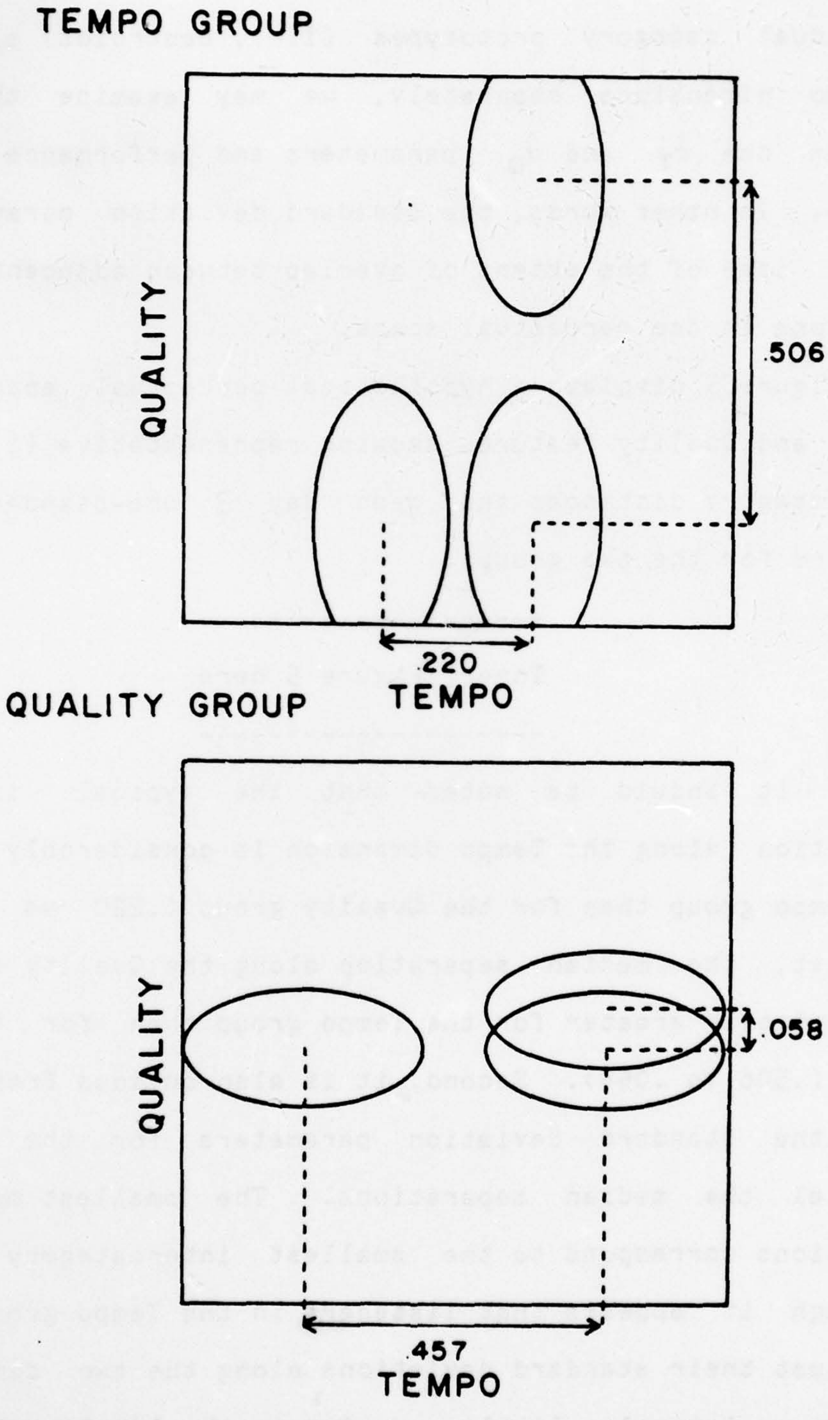

Figure 5. Hypothetical perceptual space for the Tempo and Quality features showing median inter-category distances. The ellipses represent approximate mean one-standard-deviation contours.

adequately discriminate the relatively small differences in percent attack required to achieve a high level of classification performance on this partition. Furthermore, since our earlier consideration of these parameters (cf. Table 5) revealed that $\hat{\sigma}_Q$ had largely stabilized by day 2 for the Quality group, these listeners may have approached their limit of discriminability along this dimension. In physical units (percent attack), the day 3 $\hat{\sigma}_Q$ for these listeners was approximately 8%.

When the corresponding data are considered for the Tempo group, we note that $\hat{\sigma}_T$ seems to level off at approximately .39 Hz. Other findings obtained in our laboratory suggest that this value may approach the jnd for amplitude modulation in this frequency range. Although the other study investigated modulation frequency sensitivity with a 400 Hz sawtooth carrier rather than noise, the results revealed that listeners could reliably discriminate .40 Hz differences (80% correct) in the 4 - 7 Hz modulation range (Burgy, 1978). These findings suggest that in the course of the present experiment listeners optimized their sensitivity to the more important of the two features. Whether or not they could maximize their sensitivity to both features with additional practice is an issue for further research.

A second major question of interest concerns the specific strategy that listeners employ to determine the relative emphasis to place on the two features. In the above analysis we saw that listeners in the two groups appear to focus on the feature emphasized by the category partition that they were required to learn. At first, one may wonder why the listeners don't simply

perform a similar "fine tuning" on both features. While this strategy would obviously lead to optimal performance, the observation that listeners don't do this, at least over the first three days in the task, suggests that it may be impossible for them to do so. In short, we have ignored any "cost" factors associated with the feature tuning process. The selective attentional processes hypothesized to underlie the tuning process may involve considerable effort, extensive practice or both. In other words, it appears that the listeners are constrained in the total amount of "fine tuning" that they can accomplish at any point in the task. As their familiarity with the stimuli and task increases, this overall constraint is reduced. This interpretation is consistent with recent limited-capacity views of human attentional processes (e.g., Kahneman, 1973).

With the above considerations in mind, our question becomes slightly different. Given that the listeners are constrained in the total amount of feature tuning that they can perform, how do they divide these resources between the two features? Although the decision model outlined above does not propose a specific decision criterion, other probabilistic decision models (e.g., Gerson & Goldstein, 1978) have suggested that listeners attempt to maximize the overall probability correct. Since no biasing factors were manipulated in the present study, it is possible that our listeners adopted a similar strategy to adjust the σ_T and σ_Q parameters. To investigate this possibility, an emphasis measure was determined for each feature in each condition. Since the estimated standard deviations are inversely related to

relative emphasis, emphasis measures, e_T and e_Q, were obtained from $1/\hat{\sigma}_T$ and $1/\hat{\sigma}_Q$, respectively. The theoretically optimal partition of overall emphasis across the two features was then determined for each condition in the experiment. In computing these values, the overall emphasis was estimated from the sum of e_T and e_Q. This value was taken to reflect the overall attentional effort expended by the listener at a particular point in the experiment. This overall value was then apportioned between the two features so as to maximize the average probability correct. In other words, the theoretically optimal partition of the overall emphasis on the Tempo and Quality components was determined. Table 6 displays the normalized observed and optimal emphasis parameters.

Insert Table 6 here

A comparison of the optimal and obtained values reveals a relatively close correspondence for the Tempo group, and a relatively poor overall correspondence for the Quality group. Pearson product-moment correlations between the optimal and obtained data confirm this observation, $r(23) = .93$, $r(23) = .52$ for the Tempo and Quality groups, respectively. Nonetheless, by day 3 the obtained emphasis parameters are well approximated by the optimal values for both groups, $r(7) = .98$, $r(7) = .96$ for the Tempo and Quality groups, respectively. This suggests that with experience, listeners learning the more difficult category partition (Quality group) became more likely to adopt an

Table 6. Normalized relative emphasis parameters for the two features by listener and day. Theoretically optimal values are presented in parentheses adjacent to the corresponding obtained values.

TEMPO GROUP

	DAY 1		DAY 2		DAY 3	
	TEMPO	QUALITY	TEMPO	QUALITY	TEMPO	QUALITY
MM	.72 (.64)	.28 (.36)	.71 (.70)	.29 (.30)	.73 (.69)	.27 (.31)
MG	.81 (.62)	.19 (.38)	.74 (.71)	.26 (.29)	.74 (.71)	.26 (.29)
PC	.83 (.58)	.17 (.42)	.71 (.64)	.29 (.36)	.70 (.60)	.30 (.40)
MK	.67 (.61)	.33 (.39)	.58 (.61)	.42 (.39)	.80 (.71)	.20 (.29)

QUALITY GROUP

	DAY 1		DAY 2		DAY 3	
	TEMPO	QUALITY	TEMPO	QUALITY	TEMPO	QUALITY
PH	.36 (.49)	.64 (.51)	.22 (.50)	.78 (.50)	.47 (.45)	.53 (.55)
TK	.46 (.49)	.54 (.51)	.13 (.43)	.87 (.57)	.14 (.22)	.86 (.78)
ML	.61 (.50)	.39 (.50)	.54 (.27)	.46 (.73)	.25 (.22)	.75 (.78)
MC	.56 (.22)	.44 (.78)	.18 (.44)	.82 (.56)	.22 (.37)	.78 (.63)

optimum-processor strategy.

These findings suggest a specific decision rule for the probabilistic decision model outlined above. Since listeners appear to allocate their fine tuning processes across the two features to maximize their overall probability correct, we assume that the decision processor places an unknown stimulus into the category having the highest a posteriori probability. Formally,

$$D(\underline{f}) = c^{(i)} \text{ if } Pr(c^{(i)}|\underline{f}) > Pr(c^{(j)}|\underline{f}) \text{ for all } i \neq j.$$

This interpretation is consistent with Goldstein's conclusion that listeners respond as optimum processors in determining the periodicity pitch of complex tones (Goldstein, 1973), and with a similar classification model and findings reported by Getty (Getty, Swets, Swets, & Green, in press).

III. EXPERIMENTS 3 and 4

A. INTRODUCTION

The results of Experiment 2 are consistent with the simple decision model outlined above. However, the decision model is based on a number of assumptions that require further empirical validation. Experiments 3 and 4 are designed to obtain further information relevant to these assumptions. Specifically, three assumptions of the model are considered: (1) that covariance is zero, i.e., it is assumed that the Tempo and Quality features are orthogonal, (2) that the category likelihood functions are Gaussian, and (3) that categories are represented psychologically by prototypes derived from the central tendency (centroids) of their members.

In Experiment 3 listeners were asked to classify each of 165

amplitude modulated noise patterns into one of the eight categories learned in Experiment 2. The test signals were synthesized to form a fine "grid" over the perceptual feature space (fifteen levels of modulation frequency and eleven levels of percent attack were combined factorially). Since the test signals had minimal overlap with the sixteen training signals (only two signals occurred in both sets), feedback was not provided. The outcome of this procedure was a set of labeled samples for each of the eight categories. Potential function techniques were applied to construct a probability density function from the labeled samples for each category (e.g., Murthy, 1965).

The method estimates likelihood functions by averaging a set of potential or possible functions across the labeled samples for each category. Given some point in the feature space, $\underline{x}$, the likelihood that it belongs to category $c^{(i)}$, $\Pr(\underline{x} \mid c^{(i)})$, is estimated by

$$\Pr(\underline{x} \mid c^{(i)}) = (1/M) \sum_{j=1}^{M} \gamma(\underline{x}, \underline{s}_j^{(i)}),$$

where $\left\{ \underline{s}_1^{(i)}, \underline{s}_2^{(i)}, \ldots, \underline{s}_M^{(i)} \right\}$ are the test signals that belong to category $c^{(i)}$ (i.e., that the listeners classified into this category), and $\gamma(\underline{x}, \underline{s})$ is a potential function.[6] In the present experiment, Gaussian potential functions were used to estimate the likelihood function for each category. While it is obvious that the selection of a particular potential function will influence the shape of the estimated likelihood function when the

number of labeled samples is relatively small, it should be noted that when certain conditions are met (cf. Meisel, 1972), the estimation procedure may be used to approximate any density function, given a sufficiently large number of samples. In particular, Gaussian potential functions will not always lead to Gaussian likelihood functions. For example, the likelihood functions could be multimodal or, if the labeled samples for a particular category are broadly distributed in the feature space, the resulting function may be flatter than a Gaussian. The parameters of the likelihood functions estimated in this manner will be examined in terms of the three assumptions described above.

In Experiment 4, listeners were asked to rate the pairwise similarity of all possible pairs of the eight category labels learned in Experiment 2--no sounds were presented. These subjective proximity data were decomposed into a two-dimensional "conceptual" feature space for the eight categories. The location of each category in this space will be compared with the category centroids to evaluate the third theoretical assumption.

B METHOD

1. Participants

The eight listeners who participated in the Experiment 2 served successively in Experiments 3 and 4.

2. Apparatus

Same as Experiments 1 and 2 with the addition of a video monitor for presenting the category labels in Experiment 4.

3. Stimuli

A set of 165 amplitude modulated noise signals was generated by combining factorially fifteen levels of amplitude modulation (3.5, 4.0, 4.25, 4.50, 4.75, ..., 6.75, 7.0, 7.5 Hz), and eleven levels of attack (0, 10, 20, ..., 80, 90, 100%). The noise carrier was as described in Experiment 1, and the modulation signals were sawtooth waveforms with the above characteristics. For Experiment 4, the stimuli were pairs of visually presented digits corresponding to the category labels learned in Experiment 2.

4. Procedure

Experiment 3 was conducted on two successive days immediately following the completion of Experiment 2. Each day consisted of two sessions. The first session was simply an extension of Experiment 2 where listeners classified sixteen sounds into eight categories with feedback. This session was included to insure that the listeners remembered the category partition they had learned in Experiment 2. In the second session, the listeners were told that they would hear samples of a large set of new sounds similar to those they had classified before, and that their task was to select the best category for each of these new sounds. Each of the 165 sounds was presented for 3-sec in a random order, and listeners indicated their response as in Experiment 2.

Experiment 4 was conducted on the last day of testing (i.e., the fifth day). Listeners were told that we wanted to know what they remembered about the eight categories they had learned.

They were asked to rate the similarity of each pair of signal categories. No specific instructions were provided regarding the criteria they should use in making their judgments; however, it was emphasized that their similarity ratings should be based on the sound of the categories.

C. RESULTS AND DISCUSSION

Since every listener classified each of the 165 sounds only twice, the data were analyzed by group to increase the number of category judgments for each signal. Group data were analyzed to determine if a modal category existed for each sound (i.e., a category given in at least three of the eight judgments). A single mode existed for 137 and 134 of the sounds for listeners in the Tempo and Quality groups, respectively. Column 2 of Table 7 indicates the number of signals included in each of the eight categories by this analysis.

Insert Table 7 here

A likelihood function was then estimated from these labeled samples for each category using the potential function technique outlined above. A covariance term was computed for each category to evaluate the orthogonality assumption of the decision model. These data are presented in column 3 of Table 7. The results are clear in indicating that for the stimuli investigated in the present study, the assumption of feature independence is reasonable. The mean covariance was -.003 and -.006 for the Tempo and Quality groups, respectively, at least an order of

Table 7. Summary of data from Experiment 3 (columns 1-7) and 4 (columns 8, 9) by stimulus category and group. No. indicates the number of modal stimuli in each category; Cov. the estimated covariance; $\underline{r}_T^2$ and $\underline{r}_Q^2$ the proportion of variance in the estimated likelihood functions that can be accounted for by a Gaussian for the Tempo and Quality dimensions, respectively; M_T and M_Q represent the coordinates for the category centroids in a normalized space estimated from Experiment 3; C_T and C_Q are the normalized category centroids estimated in Experiment 4.

TEMPO GROUP

CATEGORY	No.	Cov.	$\underline{r}_T^2$	$\underline{r}_Q^2$	M_T	M_Q	C_T	C_Q
1	8	.022	.982	.958	2.95	-.25	-.545	-.379
2	10	-.001	.996	.714	2.93	3.07	-.428	.347
3	12	-.002	.960	.988	3.79	.42	-.231	-.397
4	34	-.037	.970	.992	4.06	2.19	-.114	.345
5	17	.081	.947	.996	5.03	.38	.264	-.367
6	16	.000	.994	.990	5.24	1.98	.275	.348
7	21	.026	.966	.986	6.07	.02	.442	-.263
8	19	-.082	.988	.992	6.34	3.21	.336	.367

QUALITY GROUP

CATEGORY	No.	Cov.	$\underline{r}_T^2$	$\underline{r}_Q^2$	M_T	M_Q	C_T	C_Q
1	6	-.029	.982	.990	2.77	-.39	-.145	-.484
2	16	-.005	.980	.945	3.70	.51	.012	-.422
3	9	.008	.976	.945	3.52	2.46	-.541	.192
4	22	.004	.968	.978	3.54	2.64	-.482	.216
5	19	-.003	.889	.998	5.39	.20	.384	-.249
6	13	-.010	.992	.902	5.99	.40	.524	-.164
7	33	.032	.986	.994	5.08	2.05	.134	.439
8	16	-.047	.990	.980	6.20	3.00	.115	.472

magnitude smaller than the variance on either dimension.

A second question of some interest concerns the shape of the estimated likelihood functions. The Gaussian assumption of the decision model was examined by determining the proportion of variance in the likelihood function for each category that can be accounted for by a Gaussian distribution. Gaussian functions were fit to the estimated likelihood functions for each category and feature using a gradient technique with a least squares criterion. Pearson product-moment correlations were then computed between the empirically estimated and best-fitting Gaussian functions. The estimated proportion of total variance accounted for by the Gaussian (i.e., $\underline{r}^2$) is indicated in columns 4 and 5 of Table 7. It is clear from these data that the estimated likelihood functions for each category are approximately bivariate Gaussian. Although this finding is consistent with the assumptions of the present model, it must be interpreted with caution. Since a relatively small number of labeled samples were used to estimate the likelihood functions, the Parzen estimation procedure may not have converged to the true density function. Nonetheless, the findings are not inconsistent with our theoretical assumptions, and the potential function method may prove useful in future research.

Finally, the third assumption of our decision model--that each category is represented psychologically as the central tendency of its two members in the perceptual space--may be evaluated in terms of two findings. First, the means of the Gaussian likelihood functions obtained in the potential function

analysis provide an estimate of the location of each category in the perceptual space. Second, the coordinates revealed for each category in the conceptual space determined for the similarity data of Experiment 4 provide a second, independent estimate of these locations.

Consider the first estimate. Category coordinates (in a normalized space) obtained from the potential function analysis are presented for the Tempo and Quality features in columns 6 and 7 of Table 7. These coordinates correspond closely to the centroids computed from the perceptual space of Experiment 1, $\underline{r}(15) = .96$ for both dimensions.

The second estimate was obtained from a multidimensional scaling analysis of the subjective proximity data of Experiment 4. The 8 by 8 off-diagonal similarity matrix for each listener was submitted to an INDSCAL metric scaling analysis. Separate analyses were performed for the two groups. In both cases, listener ratings were well approximated by inter-stimulus distances in a two-dimensional conceptual space (the two-dimensional solution accounted for approximately 82 and 90% of the variance for the Tempo and Quality groups, respectively). Furthermore, the category coordinates revealed in this analysis (columns 8 and 9 of Table 7) correspond reasonably well to the category centroids, $\underline{r}(15) = .92$ and $\underline{r}(15) = .96$ for Tempo and Quality, respectively. These data clearly indicate that the prototype assumption is reasonable in the present experiment. Additional research would obviously be necessary to evaluate the assumption for the general case.

IV. GENERAL DISCUSSION

The primary purpose of the present study was to examine the relation between the feature extraction and decision stages in the classification of complex acoustic patterns. Several conclusions were indicated by our findings. First, the multidimensional scaling analysis of sixteen amplitude modulated noise signals presented in Experiment 1 revealed two perceptual features: Tempo--corresponding to the signal modulation frequency, and Quality--corresponding to signal attack. The results suggested that perceptual differences in signal Quality were more closely related to the percent attack (i.e., the proportion of each period spent in attack) than to the absolute duration of the attack. In other words, constant physical differences in attack become smaller perceptually as the modulation rate increases. This interpretation parallels Warren and Ackroff's (1976) finding that listeners are limited in their ability to resolve brief-duration (less than 200 msec) individual components of repeating auditory patterns. Although overall, the results of Experiment 1 were not surprising, considering the highly structured test stimuli, the analysis did provide a precise quantitative characterization of the underlying feature space.

Second, the decision model outlined above was shown to provide a reasonable fit to the classification data of Experiment 2. The model assumes that the decision process operates on the output of the feature extraction stage. Since the feature extraction process is assumed to be noisy, the decision processor

must operate in the presence of uncertainty. In the model, this uncertainty is represented by bivariate-Gaussian likelihood functions centered at the centroid for each category in the perceptual space. The decision processor simply compares the probability of each category given a particular stimulus (Equation 2) to determine its classification. An important assumption of the model is that listeners can perform a fine-tuning of the feature extraction stage to selectively increase the importance of particular features in the decision process. In the model, the effect of the tuning process is represented by a decrease in the variability of the likelihood functions. Selective tuning involves the reduction of variability along one dimension relative to another.

Both overall and selective feature tuning were observed in the present experiment. As listeners gained experience in the task, variability on both features decreased. In the model, this overall tuning accompanies the learning process where listeners reduce their overall uncertainty about the two signal parameters. As learning progresses, the listener observes that the two features are not equally important in discriminating among the eight categories. At this point selective tuning occurs to reduce the variability of the more important feature relative to the less important one.

These results are consistent with a similar attentional phenomenon observed by Watson and his associates (Watson, Kelly, & Wroton, 1976) in the discrimination of word-length tonal patterns. Each pattern consisted of a sequence of ten individual

40-msec tones. Watson et al. (1976) noted that the listeners' ability to resolve frequency differences in individual components is greatly improved when they know which component is likely to differ. In fact, under conditions of minimal uncertainty, their listeners could discriminate frequency differences in individual components of tonal sequences almost as well as they could in isolated tones. They discuss these findings in terms of a "spectral and temporal focusing of attention", and suggest that listening to complex auditory patterns may be analogous to looking at a complex picture. In the same way that viewers may focus on various aspects of a picture, listeners may attend to various aspects of a complex acoustic pattern. In both cases, knowing where to "look" for likely differences can lead to improved performance.

In the present study, listeners learned to selectively focus their attention on the more important of the two auditory dimensions. The data further suggest that selective feature tuning is not an all-or-none process since listeners did not immediately and exclusively minimize variability on the more important feature. Rather, it appears that the total amount of fine tuning that can occur is limited at any point in time. One factor that influences this limit is the amount of listener experience in the task--as listeners gain additional experience, an increased amount of fine tuning can occur. Of particular interest is the strategy that listeners use to allocate their limited attention across the two dimensions. Our data suggest that listeners employ an optimum processor strategy to determine

the extent of fine tuning to apply to the two features. In other words, they select a distribution of emphasis across the two dimensions that nearly optimizes their probability correct, given the overall limit on the amount of focusing that can occur. This conclusion is similar to that reported by Goldstein (Goldstein, 1973; Gerson & Goldstein, 1978) in his work on periodicity pitch perception.

The above results indicate that listeners have considerable flexibility in their feature extraction processes. A flexible feature extraction process of this sort can readily adapt to changing task demands. In the present study, for example, a clear difference in relative feature importance or salience was observed in the similarity judgment and classification tasks. In Experiment 1 where the data were obtained in a pairwise comparison procedure, listeners tended to emphasize signal Quality relative to Tempo (46 and 23% of the variance, respectively). Quite a different picture emerged in Experiment 2 where the listeners were trained to classify the sounds into eight categories. In this case the relative subjective importance of the two features reflected the criteria used by the experimenter to determine the eight categories. In Experiment 4 when listeners rated category similarity from memory immediately following their classification training, one might have expected the relative feature salience to parallel that observed in the classification task. However, somewhat surprisingly, the findings more closely paralleled those of Experiment 1. Listeners in both groups strongly emphasized Quality in comparing

categories from memory (28 and 55% for the Tempo group, 21 and 69% for the Quality group). It appears, then, that when comparing stimuli in a similarity judgment task, listeners tend to emphasize signal Quality relative to signal Tempo regardless of whether the signals are actually present or not. These findings clearly stress the role of task factors in determining feature saliency.

Overall, the above findings suggest that the feature extraction and decision stages interact--the decision outcome influences the feature extraction process through the hypothesized feature tuning process. Although a precise specification of the feature tuning process is not possible at this time, it is clear that any future theoretical treatment of auditory classification must adopt a more dynamic view of the feature extraction process than has been the case traditionally (cf. Howard & Ballas, 1978).

ACKNOWLEDGEMENTS

This research was sponsored by a contract from the Engineering Psychology Programs, Office of Naval Research to the Catholic University of America. James H. Howard, Jr. is the principal investigator. The authors acknowledge the helpful comments of David J. Getty , Robert P. Gurney and Darlene V. Howard, and the assistance of James A. Galgano and Mary F. Santonastasso in conducting the experiments.

REFERENCES

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. Distinctive features, categorical perception, and probability learning: Some applications of a neural model. Psychological Review, 1977, 84, 413-451.

Burgy, D. C. Hemispheric asymmetries in the perception of non-speech sound characteristics. Unpublished masters thesis, The Catholic University of America, 1978.

Carroll, J. D. & Chang, J. J. Analysis of individual differences in multidimensional scaling via an N-way generalization of 'Eckart-Young' decomposition. Psychometrika, 1970, 35, 283-319.

Cermak, G. W. & Cornillon, P. C. Multidimensional analyses of judgments about traffic noise. Journal of the Acoustical Society of America. 1976, 59, 1412-1420.

Dallos, P. The auditory periphery: Biophysics and physiology. New York: Academic Press, 1973.

Evans, E. F. & Wilson, J. P. (Eds.), Psychophysics and physiology of hearing. New York: Academic Press, 1977.

Gerson, A. & Goldstein, J. L. Evidence for a general template in central optimal processing for pitch of complex tones. Journal of the Acoustical Society of America, 1978, 63, 498-510.

Getty, D. J., Swets, J. A., Swets, J. B. & Green, D. M. On the prediction of confusion matrices from similarity judgments (Technical Report No. 3719). Cambridge, MA: Bolt, Beranek, and Newman, in press.

Goldstein, J. L. An optimum processor theory for the central formation of pitch of complex tones. Journal of the Acoustical Society of America, 1973, 54, 1496-1516.

Green, D. M. & Swets, J. A. Signal detection theory and psychophysics. New York: Wiley, 1966.

Grey, J. M. Multidimensional perceptual scaling of musical timbres. Journal of the Acoustical Society of America, 1977, 61, 1270-1277.

Howard, J. H., Jr. Psychophysical structure of eight complex underwater sounds. Journal of the Acoustical Society of America, 1977, 62, 149-156.

Howard, J. H., Jr. & Ballas, J. A. Feature selection in auditory perception (Technical Report ONR-78-5). Washington, D.C.: The Catholic University of America Human Performance Laboratory, July, 1978.

Howard, J. H., Jr. & Silverman, E. B. A multidimensional scaling analysis of 16 complex sounds. Perception & Psychophysics, 1976, 19, 193-200.

Kahneman, D. Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Klein, W., Plomp, R., & Pols, L. C. W. Vowel spectra, vowel spaces, and vowel identification. Journal of the Acoustical Society of America, 1970, 48, 999-1009.

Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika, 1964, 29, 1-27.

Meisel, W. S. Computer-oriented approaches to pattern

recognition. New York: Academic Press, 1972.

Miller, J. R. & Carterette, E. C. Perceptual space for musical structures. _Journal of the Acoustical Society of America_, 1975, _58_, 711-720.

Morgan, B. J., Woodhead, M. M., & Webster, J. C. On the recovery of physical dimensions of stimuli, using multidimensional scaling. _Journal of the Acoustical Society of America_, 1976, _60_, 186-189.

Murthy, V. K., Estimation of probability density. _Annals of Mathematical Statistics_, 1965, _35_, 1027-1031.

Plomp, R. _Aspects of tone sensation: A psychophysical study_. New York: Academic Press, 1976.

Plomp, R. & Steeneken, H. J. M. Effect of phase on the timbre of complex tones. _Journal of the Acoustical Society of America_, 1969, _46_, 409-421.

Reed, S. K. Pattern recognition and categorization. _Cognitive Psychology_, 1972, _3_, 382-407.

Shepard, R. N., Introduction to volume I. In R. N. Shepard, A. K. Romney, S. B. Nerlove (Eds.), _Multidimensional scaling: Theory and applications in the behavioral sciences_ (_Vol I_). New York: Seminar, 1972(a).

Shepard, R. N. Psychological representation of speech sounds, In E. E. David & P. B. Denes (Eds.), _Human communication: A unified view_. New York: McGraw-Hill, 1972(b).

Siebert, W. M. Stimulus transformations in the peripheral auditory system. In P. A. Kolers & M. Eden (Eds.), _Recognizing patterns: Studies in living and automatic_

systems. Cambridge, Mass.: MIT, 1968.

Takane, Y., Young, F.W., & de Leeuw, J. Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. Psychometricka, 1977,42, 7-67.

Warren, R. M. & Ackroff, J. M. Two types of auditory sequence perception. Perception & Psychophysics, 1976, 20, 387-394.

Watson, C. S., Kelly, W. J., & Wroton, H. W. Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. Journal of the Acoustical Society of America, 1976, 60, 1176-1186.

FOOTNOTES

[1]Listeners heard the stimulus pairs under each of three conditions blocked on successive days: left-ear monaural, right-ear monaural, and binaural. This factor was included for other purposes, and since a preliminary analysis revealed that data from the three presentation conditions were identical, the distinction will not be considered further.

[2]For purposes of comparison, these data were also analyzed with the ALSCAL nonmetric individual differences multidimensional scaling program (Takane, Young & de Leeuw, 1977). The resulting two-dimensional stimulus space was almost identical to that obtained in the INDSCAL analysis (Pearson product-moment correlation was $\underline{r}(15) = .999$ for both the Tempo and Quality dimensions).

[3]In the present experiment each category was made up of only two adjacent stimuli in the feature space. Consequently it is virtually impossible to distinguish the proposed prototype model from any of a number of reasonable alternatives (cf. Reed, 1972).

[4]A formally equivalent perspective would be to assume fixed signal uncertainty (i.e., the likelihood distribution for each category would have a fixed, identical standard deviation, σ, on each dimension), and a variable normalized weight vector ($\underline{w} = (\underline{w}_T, \underline{w}_Q)$, $||\underline{w}|| = 1.0$) that determines the relative importance of the two features. For this case the exponent term in Equation 1 becomes

$$-1/2(\underline{f} - \underline{p}^{(i)})\ \underline{WV}^{-1}(\underline{f} - \underline{p}^{(i)})'$$

where $\underline{W} = w\underline{I}$, and $\underline{V} = \sigma\underline{I}$. Here, the listener is assumed to adjust the weights for the two features by some selective attentional process. As the attentional weight for one feature increases relative to the other, that feature plays an increasingly important role in the decision process. From this point of view, an increase in a saliency weight corresponds to a decrease in the standard deviation parameters (i.e., σ_T, σ_Q) discussed in the text.

[5]When the a priori probabilities are equal as in this case, a decision process based on the a posteriori probabilities is equivalent to a decision process based on the likelihoods. The former approach is used here for generality.

[6]In our case the following potential function was used

$$\gamma(\underline{x},\underline{s}_j^{(i)}) = \exp(-\|\underline{x} - \underline{s}_j^{(i)}\|^2)$$

where $\|\underline{y}\|$ designates the Euclidean norm.

DIRECTOR, ENGINEERING PSYCHOLOGY
PROGRAMS, CODE 455
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY STREET
ARLINGTON, VA 22217

DEFENSE DOCUMENTATION CENTER
CAMERON STATION
ALEXANDRIA, VA 22314

DR. STEPHEN J. ANDRIOLE
ACTING DIRECTOR
CYBERNETICS TECHNOLOGY OFFICE
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD
ARLINGTON, VA 22209

CDR P. A. CHATELIER
OAD(E&LS) OUSDR&E
PENTAGON, ROOM 3D129
WASHINGTON, D.C. 20301

DIRECTOR, ACOUSTIC TECHNOLOGY
PROGRAM, CODE 222
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY STREET
ARLINGTON, VA 22217

DIRECTOR, PHYSIOLOGY PROGRAM
CODE 441
OFFICE OF NAVAL RESEARCH
800 NORTH QUINCY STREET
ARLINGTON, VA 22217

PROGRAM DIRECTOR
OCEAN TECHNOLOGY
NAVAL OCEAN RESEARCH & DEVELOPMENT
(NORDA)
BAY ST. LOUIS, MS 39529

COMMANDING OFFICER
ONR BRANCH OFFICE
ATTN: DR. J. LESTER
BUILDING 114, SECTION D
666 SUMMER STREET
BOSTON, MA 02210

COMMANDING OFFICER
ONR BRANCH OFFICE
ATTN: DR. CHARLES DAVIS
536 SOUTH CLARK STREET
CHICAGO, IL 60605

COMMANDING OFFICER
ONR BRANCH OFFICE
ATTN: DR. E. GLOYE
1030 EAST GREEN STREET
PASADENA, CA 91106

DR. BRUCE MC DONALD
OFFICE OF NAVAL RESEARCH
SCIENTIFIC LIAISON GROUP
AMERICAN EMBASSY, ROOM A-407
APO SAN FRANCISCO 96503

DIRECTOR
NAVAL RESEARCH LABORATORY
TECHNICAL INFORMATION DIVISION
CODE 2627
WASHINGTON, D.C. 20375

NAVAL RESEARCH LABORATORY
ATTN: CODE 5707
WASHINGTON, D.C. 20375

OFFICE OF THE CHIEF OF NAVAL
OPERATIONS, OP987H
PERSONNEL LOGISTICS PLANS
DEPARTMENT OF THE NAVY
WASHINGTON, D. C. 20350

DR. ANDREAS B. RECHNITZER
NAVAL POSTGRADUATE SCHOOL
OCEAN ENGINEERING
MONTERY, CA 93940

MR. ARNOLD RUBINSTEIN
NAVAL MATERIAL COMMAND
NAVMAT 08T24
DEPARTMENT OF THE NAVY
WASHINGTON, D.C. 20360

COMMANDER
NAVAL AIR SYSTEMS COMMAND
HUMAN FACTORS PROGRAMS, AIR 340F
WASHINGTON, D.C. 20361

COMMANDER
NAVAL ELECTRONICS SYSTEMS COMMAND
HUMAN FACTORS ENGINEERING BRANCH
CODE 4701
WASHINGTON, D.C. 20360

A. E. BISSON
CODE 1939
NSRDC
CARDEROCK, MD 20084

LCDR T. W. SCHROPP
NAVAL SEA SYSTEMS COMMAND
NAVSEA OOC-DA
WASHINGTON, D. C. 20362

MR. JAMES JENKINS
NAVAL SEA SYSTEMS COMMAND
CODE 06H1-3
WASHINGTON, D.C. 20362

DR. JAMES CURTIN
NAVAL SEA SYSTEMS COMMAND
PERSONNEL & TRAINING ANALYSES
NAVSEA 074C1
WASHINGTON, D.C. 20362

DIRECTOR
BEHAVIORAL SCIENCES DEPARTMENT
NAVAL MEDICAL RESEARCH INSTITUTE
BETHESDA, MD 20014

DR. GEORGE MOELLER
HUMAN FACTORS ENGINEERING BRANCH
SUBMARINE MEDICAL RESEARCH LAB.
NAVAL SUBMARINE BASE
GROTON, CT 06340

MR. PHILLIP ANDREWS
NAVAL SEA SYSTEMS COMMAND
NAVSEA 0341
WASHINGTON, D. C. 20362

BUREAU OF NAVAL PERSONNEL
SPECIAL ASSISTANT FOR RESEARCH
LIAISON
PERS-OR
WASHINGTON, D. C. 20370

NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
MANAGEMENT SUPPORT DEPARTMENT
CODE 210
SAN DIEGO, CA 92152

DR. FRED MUCKLER
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
MANNED SYSTEMS DESIGN, CODE 311
SAN DIEGO, CA 92152

MR. A. V. ANDERSON
NAVY PERSONNEL RESEARCH AND
DEVELOPMENT CENTER
CODE 302
SAN DIEGO, CA 92152

LCDR P. M. CURRAN
HUMAN FACTORS ENGINEERING BRANCH
CREW SYSTEMS DEPT., CODE 4021
NAVAL AIR DEVELOPMENT CENTER
JOHNSVILLE
WARMINSTER, PA 18950

LCDR WILLIAM MORONEY
HUMAN FACTORS ENGINEERING BRANCH
CODE 1226
PACIFIC MISSILE TEST CENTER
POINT MUGU, CA 93042

HUMAN FACTORS SECTION
SYSTEMS ENGINEERING
TEST DIRECTORATE
U.S. NAVAL AIR TEST CENTER
PATUXENT RIVER, MD 20670

DR. JOHN SILVA
MAN-SYSTEM INTERACTION DIVISION
CODE 823
NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CA 92152

HUMAN FACTORS ENGINEERING BRANCH
NAVAL SHIP RESEARCH AND
DEVELOPMENT CENTER
ANNAPOLIS DIVISION
ANNAPOLIS, MD 21402

COMMANDING OFFICER
NAVAL COASTAL SYSTEMS LAB
PANAMA CITY, FL 32401

DR. ROBERT FRENCH
NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CA 92152

DR. JERRY C. LAMB
DISPLAY BRANCH
CODE TD112
NAVAL UNDERWATER SYSTEMS CENTER
NEW LONDON CT 06320

NAVAL TRAINING EQUIPMENT CENTER
ATTN: TECHNICAL LIBRARY
ORLANDO, FL 32813

HUMAN FACTORS DEPARTMENT
CODE N215
NAVAL TRAINING EQUIPMENT CENTER
ORLANDO, FL 32813

DR. ALFRED F. SMODE
TRAINING ANALYSIS AND EVALUATION
NAVAL TRAINING EQUIPMENT CENTER
CODE N-OOT
ORLANDO, FL 32813

DR. GARY POOCK
OPERATIONS RESEARCH DEPARTMENT
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93940

DR. A. L. SLAFKOSKY
SCIENTIFIC ADVISOR
COMMANDANT OF THE MARINE CORPS
CODE RD-1
WASHINGTON, D.C. 20380

DR. JOSEPH ZEIDNER
ACTING TECHNICAL DIRECTOR
U.S. ARMY RESEARCH INST.
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

TECHNICAL DIRECTOR
U.S. ARMY HUMAN ENGINEERING LABS
ABERDEEN PROVING GROUND
ABERDEEN, MD 21005

U.S. AIR FORCE OFFICE OF
SCIENTIFIC RESEARCH
LIFE SCIENCES DIRECTORATE, NL
BOLLING AIR FORCE BASE
WASHINGTON, D.C. 20332

DR. DONALD A. TOPMILLER
SYSTEMS ENGINEERING BRANCH
HUMAN ENGINEERING DIVISION
USAF AMRL/HES
WRIGHT-PATTERSON AFB, OH 45433

LT. COL. JOSEPH A. BIRT
HUMAN ENGINEERING DIVISION
AEROSPACE MEDICAL RESEARCH LAB.
WRIGHT PATTERSON AFB., OH 45433

AIR UNIVERSITY LIBRARY
MAXWELL AIR FORCE BASE, AL 36112

DR. ROBERT WILLIGES
HUMAN FACTORS LABORATORY
VIRGINIA POLYTECHNIC INSTITUTE
130 WHITTEMORE HALL
BLACKSBURG, VA 24061

DR. ARTHUR I. SIEGEL
APPLID PSYCHOLOGICAL SERVICES
404 EAST LANCASTER STREET
WAYNE, PA 19087

DR. ROBERT R. MACKIE
HUMAN FACTORS RESEARCH, INC.
SANTA BARBARA RESEARCH PARK
6780 CORTONA DRIVE
GOLETA, CA 93017

DR. GERSHON WELTMAN
PERCEPTRONICS, INC.
6271 VARIEL AVENUE
WOODLAND HILLS, CA 91364

DR. J. A. SWETS
BOLT,BERANEK & NEWMAN, INC
50 MOULTON STREET
CAMBRIDGE, MA 02138

DR. ROSS L. PEPPER
NAVAL OCEAN SYSTEMS CENTER
HAWAII LABORATORY
P.O. BOX 997
KAILUA, HI 96734

DR. MEREDITH CRAWFORD
5606 MONTGOMERY STREET
CHEVY CHASE, MD 20015

DR. ROBERT G. PACHELLA
UNIVERSITY OF MICHIGAN
DEPT. OF PSYCHOLOGY
HUMAN PERFORMANCE CENTER
330 PACKARD ROAD
ANN ARBOR, MI 48104

DR. ROBERT FOX
VANDERBILT UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
NASHVILLE, TN 37240

DR. JESSE ORLANSKY
INSTITUTE FOR DEFENSE ANALYSES
400 ARMY-NAVY DRIVE
ARLINGTON, VA 22202

DR. STANLEY DEUTSCH
OFFICE OF LIFE SCIENCES
HQS, NASA
600 INDEPENDENCE AVE.
WASHINGTON, D.C. 20546

DR. WILLIAM A. MC CLELLAND
HUMAN RESOURCES RESEARCH OFFICE
300 N. WASHINGTON STREET
ALEXANDRIA, VA 22314

DR. WILLIAM R. UTTAL
UNIVERSITY OF MICHIGAN
INSTITUTE FOR SOCIAL RESEARCH
ANN ARBOR, MI 48106

DR. DAVID J. GETTY
BOLT, BERANEK & NEWMAN
50 MOULTON STREET
CAMBRIDGE, MA 02138

DR. RICHARD R. ROSINSKI
UNIVERSITY OF PITTSBURGH
DEPARTMENT OF INFORMATION
SCIENCE
PITTSBURGH, PA 15260

DIRECTOR, HUMAN FACTORS WING
DEFENSE & CIVIL INSTITUTE OF
ENVIRONMENTAL MEDICINE
P.O. BOX 2000
DOWNSVILLE, TORONTO, ONTARIO
CANADA

DR. A.D. BADDELEY
DIRECTOR, APPLIED PSYCH. UNIT
MEDICAL RESEARCH COUNCIL
15 CHAUCER ROAD
CAMBRIDGE, CB2 2EF
ENGLAND